



Novel mutations in NLGN3 causing autism spectrum disorder and cognitive impairment

Journal:	<i>Human Mutation</i>
Manuscript ID	humu-2019-0017.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	<p>Quartier, Angélique; Institut de Genetique et de Biologie Moleculaire et Cellulaire, neurogenetics and translational medecine courraud, jeremie; Institut de Genetique et de Biologie Moleculaire et Cellulaire, neurogenetics and translational medecine Ha, Thuong; Women's and Children's Hospital, Neurogenetics; The University of Adelaide, McGillivray, George ; Victorian Clinical Genetics Service, Royal Children's Hospital Isidor, Bertrand; Service de Génétique Médicale, CHU de Nantes Rose, Katherine; Monash Medical Centre, DROUOT, NATHALIE; IGBMC savidan, marie-armel; Institut de Genetique et de Biologie Moleculaire et Cellulaire, neurogenetics and translational medecine Jagline, Hélène; Institut de Genetique et de Biologie Moleculaire et Cellulaire, neurogenetics and translational medecine Chelly, Jamel; Institut de Genetique et de Biologie Moleculaire et Cellulaire, neurogenetics and translational medecine; Hôpitaux universitaires de Strasbourg, Laboratoire de diagnostic génétique Shaw, Marie; University of Adelaide, Paediatrics and Robinson Research Institute Laumonnier, Frédéric; INSERM U930, Genetics of autism and mental retardation; Gecz, Jozef; University of Adelaide, Medicine; The University of Adelaide and the Robinson Research Institute Mandel, Jean-Louis; IGBMC, Translational Medicine and Neurogenetics Piton, Amélie; Hôpitaux universitaires de Strasbourg, Laboratoire de diagnostic génétique; Institut de Genetique et de Biologie Moleculaire et Cellulaire, neurogenetics and translational medecine</p>
Key Words:	NLGN3, autism spectrum disorder, intellectual disability, Unfolded Protein Response

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1

2

325

4

526

6

727

8

9

1028

11

1229

13

1430

15

1631

17

18

1932

20

2133

22

2334

24

25

2635

27

2836

29

3037

31

3238

33

3439

35

3640

37

3841

39

4042

41

4243

43

44

4544

46

47

4845

49

50

5146

52

53

5447

55

56

5748

58

59

6049

The authors declare no conflict of interest.

Address for correspondence:

Amélie Piton, PhD

Laboratoire "Mécanismes génétiques des maladies neurodéveloppementales",

IGBMC

1, rue Laurent Fries,

67400, Illkirch, France

Tel: +33369551652

E-mail: piton@igbmc.fr

Number of words in the body of the manuscript: 4 140

Number of Figure: 4

Number of Tables: 2

Number of Supplementary Tables: 2

ABSTRACT

The X-linked *NLGN3* gene, encoding a postsynaptic cell adhesion molecule, was involved in a non-syndromic monogenic form of Autism Spectrum Disorder (ASD) by the description of one unique missense variant, p.Arg451Cys (Jamain et al. 2003). We investigated here the pathogenicity of additional missense variants identified in two multiplex families with intellectual disability (ID) and ASD: c.1789C>T, p.Arg597Trp, previously reported by our group (Redin et al. 2014) and present in three affected cousins and c.1540C>T, p.Pro514Ser, identified in two affected brothers. Overexpression experiments in HEK293 and HeLa cell lines revealed that both variants affect the level of the mature NLGN3 protein, its localization at the plasma membrane and its presence as a cleaved form in the extracellular environment, even more drastically than what was reported for the initial p.Arg451Cys mutation. The variants also induced an Unfolded Protein Response (UPR), probably due to the retention of immature NLGN3 proteins in the endoplasmic reticulum. In comparison, the c.1894A>G, p.Ala632Thr and c.1022T>C, p.Val341Ala variants, present in males from the general population, have no effect. Our report of two missense variants affecting the normal localization of NLGN3 in a total of five affected individuals reinforces the involvement of the *NLGN3* gene in a neurodevelopmental disorder characterized by ID and ASD.

KEYWORDS

Neurexins 3, Missense variants, Unfolded Protein Response, Autism Spectrum Disorder, Intellectual disability

BACKGROUND

Autism spectrum disorders (ASD) are characterized by deficits in social communication, restricted interests and/or stereotyped repetitive behaviors, often associated with impairment in language development and intellectual disability (ID). The role of genetics in ASD is important and if the genetic basis is at least partly multifactorial¹, a large number of "monogenic" syndromes associating ASD with ID have been identified in the last decades². One of these genes, the neuroligin-3 (*NLGN3*), was involved in ASD fifteen years ago with the identification of a unique missense variant, c.1351C>T; Arg451Cys, in two affected brothers³. This variant affects a conserved residue located in the extracellular esterase-homology domain of *NLGN3* and was shown to alter the transport of the protein and causes its retention in the endoplasmic reticulum. The greatly reduced protein fraction reaching the plasma membrane has a decreased binding affinity for its presynaptic partner *NRXN1* suggesting that the Arg451Cys variant is a loss-of-function variant^{4,5}. This mutation has also been introduced into the mouse, the knock-in (KI) animals presenting no obvious brain abnormalities, even if some subtle changes in the size of brain structure have been reported afterwards^{6,7}. An increase in inhibitory synaptic transmission was observed in cortex while an increase in excitatory synaptic transmission and long-term potentiation was detected in hippocampus⁶. As the authors did not find any alterations of inhibitory synapses in *Nlgn3* Knock-Out (KO) mice, they suggested that the Arg451Cys variant ultimately may behave as a gain-of-function mutation. Only modest impairments in social interactions were reported in this KI mice^{6,8} but not reproduced in a second KI model⁹, suggesting an influence of the genetic background in the expressivity of the mutation in mice^{10,11}. Beside this unique mutation identified in human, other arguments also supported a role of *NLGN3* in ASD as *NLGN3* expression was found to be decreased in 1) lymphoblastoid cell lines from individuals with ASD (n=35) compared to control individuals (n=35)¹², 2) some brain tissues from mice treated in utero with valproic acid (environmental rodent model of autism)¹³.

The neuroligins are a family of postsynaptic cell-adhesion molecules that are ligands for neurexins (NRXN), another class of synaptic cell-adhesion molecules, located at the presynaptic side¹⁴. These interactions promote synaptic connections between neurons. They also interact by their short intracellular domain with proteins of the post-synaptic density¹⁵. If neuroligins are not required for synapse formation *in vivo*¹⁶, they contribute to the proper synaptic functioning. NLGN3 is located both at excitatory and inhibitory synapses¹⁵. A recent study describes a novel role of NLGN3 as a secreted protein¹⁸. The authors demonstrated that the extracellular domain of neuronal NLGN3 can be cleaved. The soluble NLGN3 generated promotes high-grade glioma (HGG) proliferation by inducing PI3K-mTOR pathway activity and activating NLGN3 expression in glioma cells.

During the resequencing of several hundred candidate genes, our group recently reported a promising missense variant in *NLGN3*, Arg597Trp, identified in two cousins with ID and autistic features¹⁹ but its functional consequences were not assessed. We report here the follow-up study of this family and the identification of an additional family with two brothers carrying a missense Pro514Ser variant. We characterized the functional consequences of these missense variants, showing that they affect NLGN3 localization and function.

METHODS

Patients and identification/follow-up of missense variants in NLGN3

The genetic investigations were approved by the local Ethics Committee of the Strasbourg University Hospital (Comité Consultatif de Protection des Personnes dans la Recherche Biomédicale (CCPPRB)). For all patients, a written informed consent for genetic testing was obtained from their legal representative. For Family 1, co-segregation analysis was performed

122 in relatives of patient III.3 to test the presence of the variant c.1789C>T, Arg597Trp by
 123 amplification using (NLGN3_Foward 5'- CCCCAGGACACCAAGTTCAT-3' and
 124 NLGN3_Reverse 5'- TGGTGGTGGACGTATAGTGG-3') and Sanger sequencing (GATC,
 125 Germany) in DNA extracted from blood. For family 2, Whole Exome Sequencing (WES) was
 126 performed on the whole nuclear family (the two affected boys and their parents, as well as the
 127 unaffected sister). Briefly, blood DNAs were captured using the Roche NimbleGen SeqCap EZ
 128 exome v3 kit and sequenced on the Illumina HiSeq 2000 platform using 2x101 bp paired-end
 129 sequencing. Reads were mapped with BWA-MEM 0.7.8-r455 and variant calling was
 130 performed using GATK 3.1-1-g07a4bf8 software packages, using default parameters and
 131 following GATK Best Practices v3. Variants were annotated using an in-house annotation
 132 pipeline and filtered using population databases (dbSNP137, 1000genomes. EVS & cg69),
 133 cosegregation in the family and predictions of pathogenicity (truncating variants and missense
 134 with CADD score above 25 & predicted to be damaging by PolyPhen2 were selected). Variants
 135 were submitted to ClinVar. All the NLGN3 variants are named according to the transcript
 136 variant NM_018977.3 which encodes a protein of 828 amino acids (NP_061850.2). This is not
 137 the longest isoform (NP_851820.1, 848 amino acids, encoded by the longest transcript
 138 NM_181303.1 that includes an additional exon after exon 1) but we observed that
 139 NM_018977.3 was the most expressed in neural stem cells (data from transcriptomic studies
 140 we performed in human neuronal stem cells²⁰).

141

142 *Predictions of variants' effects*

143 All the NLGN3 variants are named according to the isoform NM_018977.3 which encodes a
 144 protein of 828 amino acids. The correspondence in the longest isoform NM_181303.1 (848
 145 amino acids) is given in Figure 1's legend. The potential effect of missense variants were
 146 predicted using SIFT (<http://sift.bii.a-star.edu.sg/>), PolyPhen2

(<http://genetics.bwh.harvard.edu/pph2/>), SNAP2 (<https://www.rostlab.org/services/SNAP/>), MutPred2 (<http://mutpred.mutdb.org/>)²¹, REVEL (<https://sites.google.com/site/revelgenomics/>), CADD (<https://cadd.gs.washington.edu/>). The presence of the different variants in the general population was checked by using ExAC database (<http://exac.broadinstitute.org/>) and later gnomAD database (<http://gnomad.broadinstitute.org/>) excluding individuals with a neurological/neurodevelopmental phenotype (“non-neuro” gnomAD). 3D modelling was predicted from protein sequence encoded by NM_018977.3 (Uniprot accession number: Q9NZ94-2) using Raptor²² and visualized using PyMol.

Characterization of the X-inactivation profile in carrier females

X-inactivation status was measured on blood DNA at the HUMARA and FRAXA loci²³. PCR amplification of both loci were performed on undigested or HpaII-digested (digestion of unmethylated allele) DNA using specific FAM-labeled primers. Amplification products were migrated on an ABI PRISM 3,500 Genetic Analyzer (Applied Biosystems, CA).

NLGN3 cloning and site-directed mutagenesis

NLGN3-001 (NM_018977.3) clone was purchased from Sino Biological Inc. (<http://www.sinobiological.com>, Cat. number: HG11160). To introduce the HA- tag (just after signal peptide) and the missense variants in *NLGN3* open reading frame (ORF), site-directed mutagenesis was performed using PFU-Turbo enzyme (18 cycles with 12min of elongation) with the following couples of primers: HA-tag forward 5'-
 tggcgctgagggccagctaccTACCCATACGATGTTCCAGATTACGCTcag
 gccccag-3' and reverse 5'-
 actgtgggtgctggggcctgAGCGTAATCTGGAACATCGTATGGGTAggtactg

172 gcc-3', Val341Ala forward 5'-aaagagtccaaggagctggCagagcaggac-3' and reverse 5'-
 173 ctggctggatgtcctgctctGccagctcctt-3', Arg451Cys forward 5'-gtgacaaccctgagacccgcTgtaaacact-
 174 3' and reverse 5'-gagtgccaccagtgttttacAgcgggtctca-3', Arg597Trp 5'-
 175 caagggtccgagatcattacTggggcactaa-3' and reverse 5'-aaaggccaccttagtgggccAgtaatgatct-3',
 176 Pro514Ser forward 5'-taccctatgttttgggggtTctatggtagg-3' and reverse 5'-
 177 gtcagtggggcctaccatagAaaccccaaaa-3' and Thr632Ala forward 5'-
 178 ccaaagtgccgcctccgatGccacccacag-3' and reverse 5'-gatgtgggagctgtgggtggCatccggaggc-3'.
 179 The PCR products were digested with DpnI and used to transform competent bacteria NEB 5-
 180 alpha Competent E. coli (NEB) by heat-shock. After bacterial amplification, plasmids were
 181 extracted using NucleoSpin® Plasmid NucleoBond® Xtra Midi (Macherey-Nagel) and
 182 sequenced (GATC, Germany) to check the presence of each variation.

184 *Cell culture and transient transfection*

185 Human HEK293 and HeLa cells were grown in DMEM supplemented by 1g/L glucose with
 186 gentamycin and respectively 10% and 5% Fetal Calf Serum in a 37°C, 5% CO₂ humidified
 187 incubator with medium renewed every two days. Cells were transfected at 60-70% of
 188 confluence in 6-well plates using Lipofectamine® 2000 DNA transfection reagent
 189 (Invitrogen®) in Opti-MEM according to manufacturer's instructions with 2µg of each *NLGN3*
 190 plasmid. Cells were stopped 24/48 hours after transfection for RNA or protein extraction. For
 191 four of the seven series of experiments, extracellular medium was also collected.

193 *Western blot and immunofluorescence*

194 For Western blot analysis, experiments were performed in replicates (4 series of HEK293 and
 195 3 serie of HeLa cells). Transfected cells were lysed in RIPA buffer with protease inhibitor

cocktail (Roche). Proteins were separated after denaturation on a 10% acrylamide gel and transferred onto a PVDF membrane. HA-tagged-NLGN3 proteins were visualized using an in-house mouse anti-HA antibody (1:1000), and their level normalized with results from beta-actin or GAPDH staining using specific antibodies (in-house for beta-actin, #MAB374, Millipore for GAPDH). To visualize the secreted form of NLGN3, extracellular media were collected and an equal amount of medium was submitted to SDS-PAGE after protein denaturation as described above. For immunofluorescence experiments, transfected cells, previously plated on coverslip in 24-wells plates, were fixed with 4% paraformaldehyde/PBS for 20 min, washed three times with PBS and incubated for 45min in blocking buffer (PBS, 10% fetal calf serum, 1% bovine serum albumin and 0.2% Triton X-100). Cells were then incubated with primary rat anti-HA antibody (1:1000, ref 11867423001, Sigma Aldrich) and secondary anti-rat antibody coupled with fluorescence. DAPI was used to stain nucleus and a mouse anti-KDEL (1:500, ab12223, abcam) to stain the endoplasmic reticulum (ER). Fluorescence was visualized on an inverted confocal microscope (SP2UV, Leica, Wetzlar, Germany).

Measure of unfolded protein response (UPR) by RT-qPCR

To measure UPR induced by the expression of WT or variant NLGN3 protein, we performed RNA extraction on several series of HEK293 (n=5) and of HeLa cells (n=5) 24h after plasmid transfection, using the RNeasy mini kit (Qiagen, Valencia, CA, USA) including a DNase treatment. 500ng to 1µg of total RNA was reverse transcribed into cDNA using random hexamers and SuperScript IV reverse transcriptase according to manufacturer's recommendation. Real-time PCR (qPCR) were performed on LightCycler 480 II (Roche) using the QuantiTect SYBR Green PCR Master Mix (Qiagen) and primers to amplify CHOP (CHOP_Foward 5'- GACCTGCAAGAGGTCCTGTC-3' and CHOP_Reverse 5'-

220 CTCCTCCTCAGT CAGCCAAG-3'), spliced isoform of XPB1 (XPB1_spliced_Reverse 5'-
 221 GCCTGCACCTGCTGCGGA-3'), ATF4 (ATF4_Foward 5'-
 222 CCAACAACAGCAAGGAGGAT-3' and ATF4_Reverse 5'-
 223 AGGTCATCTGGCATGGTTTC-3') and BiP (BiP_Foward 5'-
 224 TGTTC AACCAATTATCAGCAA ACTC-3' and BiP_Reverse 5'-
 225 TTCTGCTGTATCCTCTTCACCAGT-3'). All qPCR reactions were performed in triplicate.
 226 Reaction specificity was controlled by post-amplification melting curve analysis. The relative
 227 expression of gene-of-interest vs *GAPDH* and *YWHAZ* was calculated using the 2-($\Delta\Delta C_t$)
 228 method. Expression of the spliced form of *XPB1*, *CHOP* and *BiP* induced by overexpression
 229 of NLGN3 variants were compared to what is obtained for overexpression of the wild-type
 230 NLGN3. To avoid effect due to the overexpression of the protein itself, we normalized the
 231 results with *NLGN3* mRNA expression instead of *GAPDH/YWHAZ*.

233 *Statistical analyses*

234 For each statistical analysis, normal distribution was first checked using Shapiro-Wilk and/or
 235 Kolmogorov-Smirnov tests. When normality is demonstrated, a t-test is performed for
 236 comparison of two means, with Welch's correction to take into account the inequality of
 237 variances. For comparisons of more than two groups, Brown-Forsythe and Welch ANOVA tests
 238 were performed in case of normal distribution to take into account the inequality of variances.
 239 If ANOVA is significant, Dunnett's T 3 multiple comparisons test were performed to compare
 240 each NLGN3 variant to the wild-type. When the data do not follow a normal distribution, a
 241 Kruskal-Wallis' ANOVA test was performed instead. If ANOVA is significant, Dunn's
 242 multiple comparison test was performed to compare each NLGN3 variant to the wild-type.

243

244

245

246 RESULTS

247 Missense variations identified in *NLGN3* in individuals with ID and ASD or autistic 248 features

249 The first missense variant Arg451Cys was identified in two Swedish brothers with ASD, one
250 with typical autism and ID and one with Asperger syndrome³. Following the discovery of this
251 mutation, the coding regions of *NLGN3* were extensively sequenced afterwards (in more than
252 1,500 probands with ASD in total, see **Table 1**). These studies screened cohorts of very different
253 geographical origins but failed to identify promising non-synonymous variant in *NLGN3*^{24–36}.
254 The screening of a Chinese cohort identified a novel missense variant Gly426Ser³⁷, but
255 functional studies performed recently could not identify any functional consequences on
256 *NLGN3* function³⁸. Another missense variant was identified by whole exome sequencing
257 (WES) performed in a cohort of simplex and multiplex families with ASD, in one boy with ID,
258 hyperactivity and dysmorphic features (c.1022T>C, Val341Ala), but the authors did not reach
259 a conclusion about its pathogenicity as no functional validation was performed³⁹. This variant
260 was afterward found in the hemizygous state in several boys from the Exome Aggregation
261 Consortium project (ExAC) and its updated version gnomAD, which was not in favor of its
262 pathogenicity. We previously identified and reported a missense variant in *NLGN3*, c.1789C>T,
263 Arg597Trp present at the hemizygote state in one boy with ID and ASD and in his maternal
264 cousin with similar phenotype¹⁹. We performed a follow-up segregation analysis in this family
265 and found that this variant was present in a third boy, III.3, cousin of III.1 and brother of III.2,
266 while it was absent from an unaffected uncle II.2 (**Figure 1A, Table 2**). All three cousins
267 present with ID, ASD or autistic traits, sleep disorder, but no dysmorphism or other anomalies,

except a mild microcephaly which was reported in the two cousins III.2 and III.3. Both presented with mild ID while Individual III.1 presented with severe ID with no speech acquisition at 6 years of age. No bias of X-inactivation status was observed in the mother II.3 and the aunt II.5, and only a potential slight bias (FRAXA: 76/24; HUMARA: 80/20) has been identified for the grand-mother I.1. In a second family, composed of two brothers affected by non-syndromic ID, autistic traits and language impairment, a family-WES analysis identified a missense variant c.1540C>T, Pro514Ser in *NLGN3*, present in both brothers and inherited from the mother. The two brothers of this second family had normal early infant development with language development until 18 months of age and then presented with significant speech regression. They subsequently had a severe disability with autistic traits and no neuropsychology assessment was possible.

279

***NLGN3* is relatively intolerant to missense changes**

These two missense variants, like the originally published pathogenic variant c.1351C>T, Arg451Cys³, affect highly conserved amino acids and are predicted to be deleterious by all the different programs (**Table 2, Figure 1B**). They all alter amino acids located in the extracellular domain of *NLGN3*. The change of the arginine to a tryptophan in position 597 and of a proline to a serine in position 514 modify the hydrogen bonds with adjacent amino acids (**Figure 1D**) and might therefore affect *NLGN3* protein structure. We reviewed all the nonsynonymous variants present in the general population using the gnomAD database (v2.1). There is no loss-of-function variant in these general populations, if we except a putative splice variant affecting a non-canonical exon specific to the longest isoform but falling in intronic region in NM_018977.3 (c.457+639G>A). A putative truncating variant falling in this exon was also reported previously during the resequencing of X-Linked ID families (c.457+666C>T in the

NM_018977.3 transcript)⁴⁰. Two and half times less missense variants was observed compared to what is expected (151/376), suggesting that *NLGN3* gene is relatively intolerant to missense changes (Z-score = 4.21)(this metrics is available in the gene constraint section in gnomAD v2.1). In total, 139 nonsynonymous changes (136 missense variants and 2 in-frame indels) are reported in individuals from the gnomAD cohorts excluding patients with neurodevelopmental or neurological conditions (gnomAD “non-neuro” cohorts) (**Supp. Table S1, Figure 2A**). Half of them only (71/139) are present at the hemizygous state in male individuals (9 only in three or more males). Prediction of deleterious effect of these missense variants, based on CADD score, shows that they range from very low effect (CADD=7) to very strong effect (CADD=34). It appears therefore difficult to discriminate, among the variants identified in patients with neurodevelopmental conditions, those being truly pathogenic only based on *in silico* predictions and without performing additional functional studies.

The two novel missense mutations affect the level of mature NLGN3 proteins

We decided to undertake a study of the effect of the two missense variants Pro514Ser and Arg597Trp on NLGN3 function, to determine their pathogenicity, compared to the effect of the unique pathogenic mutation validated to date Arg451Cys (**Table 2**). We also included two rare missense variants reported in males from gnomAD cohorts, as negative controls, Val341Ala (first selected because it was observed in a patient with ASD⁴¹ but finally reported in 7 males from “non-neuro” gnomAD cohorts) and Thr632Ala (reported in 138 males from “non-neuro” gnomAD cohorts and found in one man from a cohort of control individuals previously sequenced²⁸)(**Figure 2A**). NLGN3 expression level was analyzed in HeLa and HEK293 cell lines. A drastic decrease in the level of NLGN3, significant for Pro514Ser and Arg597Trp but

not for Arg451Cys, was observed for variant proteins. This decrease concerns especially the mature glycosylated protein (upper band). On the contrary, the Val341Ala and Thr632Ala variants lead to an expression level similar to wild-type (**Figure 2B, upper panel**).

Missense variants affect NLGN3 membrane localization and its secretion

By immunostaining, we observed that wild-type NLGN3 and variant NLGN3 carrying the variants Val341Ala or Thr632Ala are localized both at the plasma membrane and in the cytoplasm when overexpressed in HEK293 (**Figure 3**) and HeLa cells (data not shown). On the contrary, we did not detect any membrane localization for the variant NLGN3 with Pro514Ser and Arg597Trp variants. For the Arg451Cys variant, a weak staining was observed at the membrane, but reduced when compared to the wild-type protein. This suggests that part of the NLGN3 proteins carrying the variant Arg451Cys are not properly addressed to the membrane, consistent with previous report⁵. Furthermore, the Arg451Cys variant was found to affect the correct folding of the extracellular domain leading to a partial retention of the protein in the endoplasmic reticulum (ER), decreasing the trafficking of NLGN3 to the cell surface^{5,42}. We confirmed this retention in ER for NLGN3 proteins with Pro514Ser and Arg597Trp variants by performing a costaining with KDEL, a marker of ER (**Figure 3**). It was recently shown that the extracellular domain of neuronal NLGN3 is cleaved and secreted into the extracellular medium¹⁸. We confirmed that we can detect a secreted NLGN3 protein in the extracellular medium of HeLa and HEK293 cells transfected with wild-type, Val341Ala or Thr632Ala *NLGN3* cDNA (**Figure 2B, lower panel**). Consistent with the absence of membrane localization, we could not detect secreted forms of the NLGN3 proteins carrying Pro514Ser and Arg597Trp variants, and only a residual amount of secreted protein was detected for the Arg451Cys variant.

340

341 **Activation of the UPR response**

342 It was previously shown that the Arg451Cys NLGN3 mutation is partially retained in the
343 endoplasmic reticulum (ER) subsequently causing ER stress and leading to activation of the
344 UPR (unfolded protein response) cell response⁴³. This stress was presumed to participate in the
345 pathogenicity of the mutation⁴⁴. We therefore tested whether some UPR markers are induced
346 after overexpression of wild-type and variant NLGN3 proteins. mRNA levels for the spliced
347 transcripts of XBP1 and for CHOP and BiP, markers of the UPR stress, are induced after a
348 treatment with thapsigargin, a drug inhibiting the ER calcium pump, in HEK293 cells (**Figure**
349 **4A**). As overexpression of the *NLGN3* plasmids, including wild-type one, lead to a slight
350 increase of UPR targets (data not shown), we normalized the UPR response with the level of
351 *NLGN3* expression (**Figure 4B**). We observed that Arg451Cys, Pro514Ser and Arg597Trp
352 missense variants significantly induced the expression of genes of the UPR in HEK293 cells
353 (**Figure 4C**). The UPR response was more pronounced for the two novel mutations as compared
354 to the initial Arg451Cys mutation, which is consistent with the fact that the expression of the
355 mature NLGN3 protein and the protein localization are more altered by the Pro514Ser and
356 Arg597Trp changes than by the Arg451Cys change (**Figure 2B and Figure 3**).

357

358

359 **DISCUSSION**

360 The *NLGN3* and *NLGN4X* genes were the first gene described as involved in a non-syndromic
361 monogenic form of ASD³. Mutations in *NLGN4X* were thereafter found in individuals with
362 XLID with or without ASD⁴⁵, however the role of *NLGN3* in NDD was still restricted in ASD
363 with the Arg451Cys mutation until now. Extensive functional studies have been performed, in

cellular and animal models, all based on the single Arg451Cys missense mutation reported. In the meantime, the coding regions of *NLGN3* were screened for mutation in numerous cohorts of ASD individuals, without finding additional convincing pathogenic variants. Systematic trio exome sequencing of ASD cohorts was also deceptive, as only two de novo missense variants are catalogued in the denovodb database, from the study by Iossifov et al. 2004⁴⁶: Arg195Trp, with a low CADD score (18.3) in a patient with ASD, and Arg797Gln, CADD = 20.7, in an unaffected sib. With two novel missense variants affecting the membrane localization of *NLGN3* and the expression of the mature glycosylated form of the protein, identified in a total of five affected individuals from two multiplex families and cosegregating with affected status in males, we confirm that missense mutations in *NLGN3* can indeed be responsible for a neurodevelopmental disorder. All seven individuals reported now with a pathogenic variant in *NLGN3* have had autistic traits. The severity of cognitive impairment is variable in patients, ranging from no ID to severe ID, even in individuals carrying the same mutation, as similarly found for *NLGN4X* mutations. Sleep disorders have been described in the three cousins of family 1, and interestingly, sleep alterations were observed in the rat model *Nlgn3* KO, with an increase of rapid eye movement (REM) and a decrease in non-rapid eye movement (NREM) sleep periods as compared to wild-type⁴⁷.

We demonstrated here that *NLGN3* proteins carrying Pro514Ser and Arg597Trp variants do not reach the plasma membrane in the cells we used for overexpression studies. This model indicates that *NLGN3* proteins with these variants would be prevented from interacting with neurexin proteins in the human brain. An induction of the UPR response, illustrated by an increase of the spliced isoform of *XPB1* and an increase of expression of *CHOP* and *BiP*, was observed in HEK293 cells. The activation of the UPR stress and the loss of the mature form of *NLGN3* strongly suggest that variant proteins are stacked in the endoplasmic reticulum (ER), which was confirmed by co-staining experiments with an ER marker, KDEL.

At this stage, we can only speculate about the underlying mechanism of action of these variants. On aggregate we found that the variants we studied: 1) prevent NLGN3 to reach the cell membrane which may result in loss of interaction between the pre- and post-synaptic compartments; that 2) they are not secreted which may have an effect on glial cells proliferation; and that 3) they induce ER stress and resulting cellular response (UPR). The two first observations suggest that these variants would have a loss-of-function effect whereas the third one (activation of UPR) would rather evoke a gain of function effect. The interpretation of this latter result is limited by the fact that this UPR stress is observed in vitro in a context of overexpression. However, a link between NLGN3 missense variants and UPR stress is supported by a recent published observation showing that alterations of synaptic function found in the Purkinje cells of the Arg451Cys KI mice (increased frequency of miniature excitatory currents) are rescued by inhibiting the UPR response ⁵⁰.

Up to date, the three pathogenic mutations described in NLGN3 are missense changes. Interpretation of missense variants in NLGN3 remains challenging. Indeed, some missense variants found in males in the general population (gnomAD non neuro cohorts) or having no observed effect on protein function³⁷ are located in the vicinity of the three pathogenic mutations (**Supp. Table S1**). *In silico* prediction tools are not sufficient to discriminate between the benign and pathogenic variants. At the time of writing, 12 missense variants were reported in Clinvar. If we except the initial Arg451Cys pathogenic variant reported by Jamain et al. as “risk factor”, our Arg597Trp variant (that we previously reported it as “likely pathogenic”) and the Thr632Ala (reported as “Benign”/“Likely benign”), all the 9 remaining variants are labeled as of “Uncertain significance” (see **Supp. Table S2**). The fact that the *NGLN3* gene is located on the X chromosome is a bias against detecting de novo events (that are more likely to occur one generation above, from the maternal grandfather) and reporting missense variants in

publications of large exome or genome cohort studies. Testing the X-inactivation status of mothers does not seem to be informative. Segregation analysis, including testing of maternal grand-parents and affected or unaffected males (brothers, maternal uncles) might be informative but are not always feasible, especially in large cohort studies. Therefore, functional testing will remain mandatory to conclude about pathogenicity of missense variants for diagnostic purposes, until more is known about structure/function relations in this transmembrane protein.

421

422 CONCLUSIONS

423 In conclusion, our study confirms that several missense variants in NLGN3 are involved in a
424 non-syndromic form of ID associated with autistic manifestations. The three pathogenic
425 variants studied lead to an absence/decrease of membrane localization of NLGN3, a retention
426 in the endoplasmic reticulum (ER) and an induction of a cellular response related to ER stress.

427

428 ACKNOWLEDGMENTS

429 The authors thank the families for their participation in this study. They thanks Didier Devys
430 form interpreting X inactivation studies, and Paola Rossolillo, from the IGBMC molecular
431 biology platform. This work was funded by Association APLM. This study was also supported
432 by grant ANR-10-LABX-0030-INRT, a French state fund managed by the Agence Nationale
433 de la Recherche under the frame program Investissements d'Avenir ANR-10-IDEX-0002-02,
434 Australian NHMRC grants APP1091593 and APP1041920 to Jozef Gecz and grants from
435 Fondation de France (2012-33662) and "Association pour le Développement de la
436 Neurogénétique" to Frédéric Laumonnier.

437

438

439 **REFERENCES**

- 440 1. Sandin S, Lichtenstein P, Kuja-Halkola R, et al. The Heritability of Autism Spectrum Disorder.
441 JAMA 2017;318(12):1182–1184.
- 442 2. Gonzalez-Mantilla AJ, Moreno-De-Luca A, Ledbetter DH, Martin CL. A Cross-Disorder Method to
443 Identify Novel Candidate Genes for Developmental Brain Disorders. JAMA Psychiatry
444 2016;73(3):275–283.
- 445 3. Jamain S, Quach H, Betancur C, et al. Mutations of the X-linked genes encoding neuroligins
446 NLGN3 and NLGN4 are associated with autism. Nat. Genet. 2003;34(1):27–29.
- 447 4. Comoletti D, De Jaco A, Jennings LL, et al. The Arg451Cys-neuroligin-3 mutation associated with
448 autism reveals a defect in protein processing. J. Neurosci. 2004;24(20):4889–4893.
- 449 5. Chih B, Afridi SK, Clark L, Scheiffele P. Disorder-associated mutations lead to functional
450 inactivation of neuroligins. Hum. Mol. Genet. 2004;13(14):1471–1477.
- 451 6. Tabuchi K, Blundell J, Etherton MR, et al. A neuroligin-3 mutation implicated in autism increases
452 inhibitory synaptic transmission in mice. Science 2007;318(5847):71–76.
- 453 7. Ellegood J, Lerch JP, Henkelman RM. Brain abnormalities in a Neuroligin3 R451C knockin mouse
454 model associated with autism. Autism Res 2011;4(5):368–376.
- 455 8. Etherton M, Földy C, Sharma M, et al. Autism-linked neuroligin-3 R451C mutation differentially
456 alters hippocampal and cortical synaptic function. Proc. Natl. Acad. Sci. U.S.A.
457 2011;108(33):13764–13769.
- 458 9. Chadman KK, Gong S, Scattoni ML, et al. Minimal aberrant behavioral phenotypes of neuroligin-
459 3 R451C knockin mice. Autism Res 2008;1(3):147–158.
- 460 10. Jaramillo TC, Liu S, Pettersen A, et al. Autism-related neuroligin-3 mutation alters social
461 behavior and spatial learning. Autism Res 2014;7(2):264–272.
- 462 11. Jaramillo TC, Escamilla CO, Liu S, et al. Genetic background effects in Neuroligin-3 mutant mice:
463 Minimal behavioral abnormalities on C57 background. Autism Res 2018;11(2):234–244.
- 464 12. Yasuda Y, Hashimoto R, Yamamori H, et al. Gene expression analysis in lymphoblasts derived
465 from patients with autism spectrum disorder. Mol Autism 2011;2(1):9.
- 466 13. Kolozsi E, Mackenzie RN, Roulet FI, et al. Prenatal exposure to valproic acid leads to reduced
467 expression of synaptic adhesion molecule neuroligin 3 in mice. Neuroscience
468 2009;163(4):1201–1210.
- 469 14. Yamagata M, Sanes JR, Weiner JA. Synaptic adhesion molecules. Curr. Opin. Cell Biol.
470 2003;15(5):621–632.

- 471 15. Chubykin AA, Liu X, Comoletti D, et al. Dissection of synapse induction by neuroligins: effect of a
472 neuroligin mutation associated with autism. *J. Biol. Chem.* 2005;280(23):22365–22374.
- 473 16. Varoqueaux F, Aramuni G, Rawson RL, et al. Neuroligins determine synapse maturation and
474 function. *Neuron* 2006;51(6):741–754.
- 475 17. Graf ER, Zhang X, Jin S-X, et al. Neurexins induce differentiation of GABA and glutamate
476 postsynaptic specializations via neuroligins. *Cell* 2004;119(7):1013–1026.
- 477 18. Venkatesh HS, Johung TB, Caretti V, et al. Neuronal Activity Promotes Glioma Growth through
478 Neuroligin-3 Secretion. *Cell* 2015;161(4):803–816.
- 479 19. Redin C, Gérard B, Lauer J, et al. Efficient strategy for the molecular diagnosis of intellectual
480 disability using targeted high-throughput sequencing. *J. Med. Genet.* 2014;51(11):724–736.
- 481 20. Quartier A, Chatrousse L, Redin C, et al. Genes and Pathways Regulated by Androgens in Human
482 Neural Cells, Potential Candidates for the Male Excess in Autism Spectrum Disorder. *Biol.*
483 *Psychiatry* 2018;
- 484 21. Pejaver V, Mooney SD, Radivojac P. Missense variant pathogenicity predictors generalize well
485 across a range of function-specific prediction challenges. *Hum. Mutat.* 2017;38(9):1092–1108.
- 486 22. Källberg M, Wang H, Wang S, et al. Template-based protein structure modeling using the
487 RaptorX web server. *Nat Protoc* 2012;7(8):1511–1522.
- 488 23. Allen RC, Zoghbi HY, Moseley AB, et al. Methylation of HpaII and HhaI sites near the
489 polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome
490 inactivation. *Am. J. Hum. Genet.* 1992;51(6):1229–1239.
- 491 24. Talebizadeh Z, Bittel DC, Veatch OJ, et al. Do known mutations in neuroligin genes (NLGN3 and
492 NLGN4) cause autism? *J Autism Dev Disord* 2004;34(6):735–736.
- 493 25. Vincent JB, Kolozsvari D, Roberts WS, et al. Mutation screening of X-chromosomal neuroligin
494 genes: no mutations in 196 autism probands. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*
495 2004;129B(1):82–84.
- 496 26. Ylisaukko-oja T, Rehnström K, Auranen M, et al. Analysis of four neuroligin genes as candidates
497 for autism. *Eur. J. Hum. Genet.* 2005;13(12):1285–1292.
- 498 27. Gauthier J, Bonnel A, St-Onge J, et al. NLGN3/NLGN4 gene mutations are not responsible for
499 autism in the Quebec population. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*
500 2005;132B(1):74–75.
- 501 28. Blasi F, Bacchelli E, Pesaresi G, et al. Absence of coding mutations in the X-linked genes
502 neuroligin 3 and neuroligin 4 in individuals with autism from the IMGSAC collection. *Am. J.*
503 *Med. Genet. B Neuropsychiatr. Genet.* 2006;141B(3):220–221.
- 504 29. Wermter A-K, Kamp-Becker I, Strauch K, et al. No evidence for involvement of genetic variants
505 in the X-linked neuroligin genes NLGN3 and NLGN4X in probands with autism spectrum
506 disorder on high functioning level. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*
507 2008;147B(4):535–537.

- 508 30. Pampanos A, Volaki K, Kanavakis E, et al. A substitution involving the NLGN4 gene associated
509 with autistic behavior in the Greek population. *Genet Test Mol Biomarkers* 2009;13(5):611–
510 615.
- 511 31. Yanagi K, Kaname T, Wakui K, et al. Identification of Four Novel Synonymous Substitutions in
512 the X-Linked Genes Neuroligin 3 and Neuroligin 4X in Japanese Patients with Autistic Spectrum
513 Disorder. *Autism Res Treat* 2012;2012:724072.
- 514 32. Avdjieva-Tzavella DM, Todorov TP, Todorova AP, et al. Analysis of the genes encoding
515 neuroligins NLGN3 and NLGN4 in Bulgarian patients with autism. *Genet. Couns.*
516 2012;23(4):505–511.
- 517 33. Steinberg KM, Ramachandran D, Patel VC, et al. Identification of rare X-linked neuroligin
518 variants by massively parallel sequencing in males with autism spectrum disorder. *Mol Autism*
519 2012;3(1):8.
- 520 34. Volaki K, Pampanos A, Kitsiou-Tzeli S, et al. Mutation screening in the Greek population and
521 evaluation of NLGN3 and NLGN4X genes causal factors for autism. *Psychiatr. Genet.*
522 2013;23(5):198–203.
- 523 35. Xu X, Xiong Z, Zhang L, et al. Variations analysis of NLGN3 and NLGN4X gene in Chinese autism
524 patients. *Mol. Biol. Rep.* 2014;41(6):4133–4140.
- 525 36. Mikhailov A, Fennell A, Plong-on O, et al. Screening of NLGN3 and NLGN4X genes in Thai
526 children with autism spectrum disorder. *Psychiatr. Genet.* 2014;24(1):42–43.
- 527 37. Xu X, Xiong Z, Zhang L, et al. Variations analysis of NLGN3 and NLGN4X gene in Chinese autism
528 patients. *Mol. Biol. Rep.* 2014;41(6):4133–4140.
- 529 38. Xu X, Hu Z, Zhang L, et al. Not all neuroligin 3 and 4X missense variants lead to significant
530 functional inactivation. *Brain Behav* 2017;7(9):e00793.
- 531 39. Yu, Chahrour, Coulter, et al. Using whole-exome sequencing to identify inherited causes of
532 autism. *Neuron* 2013;77(2):259–273.
- 533 40. Hu H, Haas SA, Chelly J, et al. X-exome sequencing of 405 unresolved families identifies seven
534 novel intellectual disability genes. *Mol. Psychiatry* 2016;21(1):133–148.
- 535 41. Yu TW, Chahrour MH, Coulter ME, et al. Using whole-exome sequencing to identify inherited
536 causes of autism. *Neuron* 2013;77(2):259–273.
- 537 42. De Jaco A, Dubi N, Comoletti D, Taylor P. Folding anomalies of neuroligin3 caused by a mutation
538 in the alpha/beta-hydrolase fold domain. *Chem. Biol. Interact.* 2010;187(1–3):56–58.
- 539 43. Ulbrich L, Favaloro FL, Trobiani L, et al. Autism-associated R451C mutation in neuroligin3 leads
540 to activation of the unfolded protein response in a PC12 Tet-On inducible system. *Biochem. J.*
541 2016;473(4):423–434.
- 542 44. Autism spectrum disorder is related to endoplasmic reticulum stress induced by mutations in
543 the synaptic cell adhesion molecule, CADM1. - PubMed - NCBI [Internet]. [date unknown];[cited
544 2018 Jul 4] Available from: <https://www.ncbi.nlm.nih.gov/pubmed/21364653>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

545 45. Laumonnier F, Bonnet-Brilhault F, Gomot M, et al. X-linked mental retardation and autism are
546 associated with a mutation in the NLGN4 gene, a member of the neuroligin family. Am. J. Hum.
547 Genet. 2004;74(3):552–557.

548 46. Iossifov I, O’Roak BJ, Sanders SJ, et al. The contribution of de novo coding mutations to autism
549 spectrum disorder. Nature 2014;515(7526):216–221.

550 47. Thomas AM, Schwartz MD, Saxe MD, Kilduff TS. Sleep/Wake Physiology and Quantitative
551 Electroencephalogram Analysis of the Neuroligin-3 Knockout Rat Model of Autism Spectrum
552 Disorder. Sleep 2017;40(10)

553

554

FIGURE LEGENDS

Figure 1. Cosegregating missense variants identified in *NLGN3* in two multiplex families with ID and autism

(A) Pedigrees of two families carrying the missense variants c.1789C>T, p.Arg597Trp and c.1540C>T, p.Pro514Ser. Open circles represent females, open squares represent unaffected males, closed squares represent affected males, wt = wild-type, * = mutation carrier, NT: not tested. The patient previously described in Redin et al.¹⁹ is the Individual III.3 **(B)** Evolutionary conservation of amino acids at the different positions affected by the missense changes identified (451, 514 and 597) across members of the human neuroligin family. **(C-D)** View of predicted 3D structure of NLGN3, with the predicted structural effects of the two missense variants Pro514Ser and Arg597Trp. Side chain changes are predicted to affect putative polar contacts (hydrogen bounds) with adjacent amino acids (yellow dot lines).

Figure 2. Schematic representation of NLGN3 protein and expression of HA-tagged wild-type and variant NLGN3 proteins in HEK293 and HeLa cells

(A) Schematic representation of the HA-tagged NLGN3 protein (encoded by the NM_018977.3 transcript) with its different extracellular, transmembrane and intracellular domains. Missense variants reported at the hemizygote state in males in “non-neuro” gnomAD populations are indicated by grey circles: variants present in one or two males are indicated by small circles while larger circles indicate variants present in three or more males. The different missense changes studied here are indicated: the initial Arg451Cys pathogenic variant reported by Jamain et al. (in red); the two variants Arg597Trp and Pro514Ser (in orange); the variant Val341Ala initially reported in an individual with ASD by Yu et al.⁴¹ but also present in 7 males from the “non-neuro” gnomAD general population cohorts and the variant Thr632Ala present in 138 males and classified as certainly benign (in green) (B) Expression of NLGN3 proteins in HEK293 and HeLa cells transiently transfected with NLGN3 constructs was detected by SDS-PAGE and immunoblotting using an anti-HA antibody, revealing the immature form (lower band) and the mature glycosylated form (higher band). Quantifications were performed using expression of reference proteins (GAPDH or actin) on a total of n=7 series of cells (n=4 HEK293 and n=3 HeLa cells). Expression of secreted cleaved form of NLGN3 was analyzed in the same manner on extracellular cell media for four of these series of cells. ANOVA Kruskal-Wallis tests and Dunn’s multiple comparisons test were performed in order to compare expression of variant NLGN3 to wild type protein: ns: not significant; * p<0.05; ** p<0.01; error bars represent SEM.

Figure 3. Cellular localization of HA-tagged wild-type and variant NLGN3 proteins in HEK293 cells

HEK293 cell lines were transiently transfected with NLGN3 constructs, HA-tagged wild-type (WT) or variant NLGN3.) Immunofluorescence experiments using an anti-HA antibody revealed the cellular localization of WT and variant NLGN3 proteins (green fluorescence) in HEK293 cells. The DAPI staining indicates the position of the nuclei and the KDEL staining shows the endoplasmic reticulum (ER)

Figure 4. Expression of markers of the Unfolded Protein Response (UPR) response in HEK293 cells.

(A) Expression of CHOP, BiP and spliced XPB1 mRNA (normalized by the expression of two reference genes, GAPDH and YWHAZ) are induced in HEK293 cells after a treatment with thapsigargin (5 hours, 1μM, n=3). t test with Welsch's correction, ** p<0.001. Error bars represent SEM (B) Increase of NLGN3 mRNA after a transient transfection of HEK293 cells with plasmids containing HA-tagged wild-type (WT) or variant NLGN3 proteins compared to Lipofectamine only (Lipo) (n=5 for each condition) (results normalized by GAPDH and YWHAZ expression). Kruskal-Wallis' ANOVA test did not revealed significant difference between the conditions; error bars represent SEM. (C) Transfection of the variant Arg451Cys, Pro514Ser and Arg597Trp NLGN3 proteins in HEK293 cells lead to increased mRNA expression level of the UPR markers CHOP, BiP and spliced transcript of XPB1 (results normalized by NLGN3 mRNA expression) compared to cells transfected with WT NLGN3. As Brown-Forsythe and Welsch's ANOVA revealed significant difference, and multiple

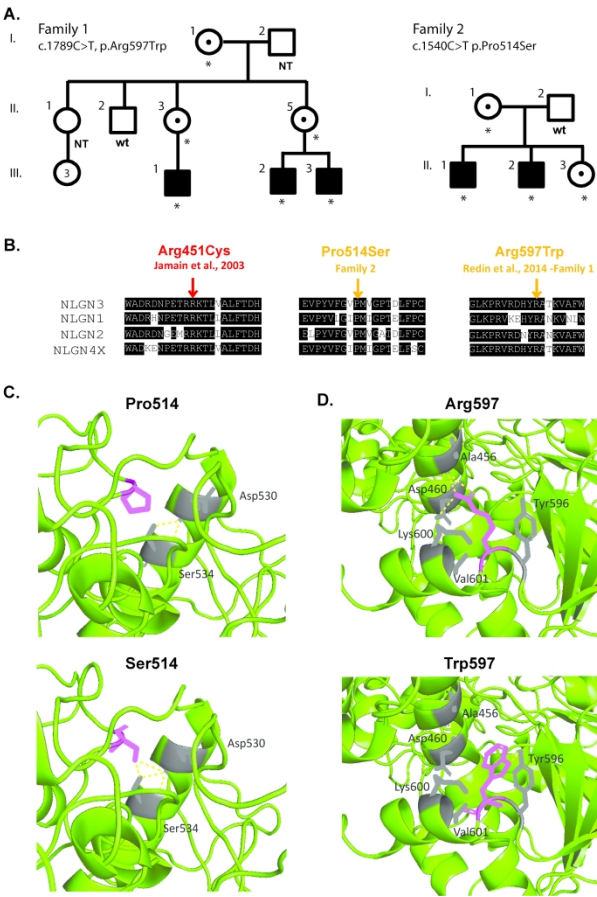
comparisons tests were performed using Dunnet's T3 test: ns: not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; error bars represent SEM.

SUPPORTING INFORMATION

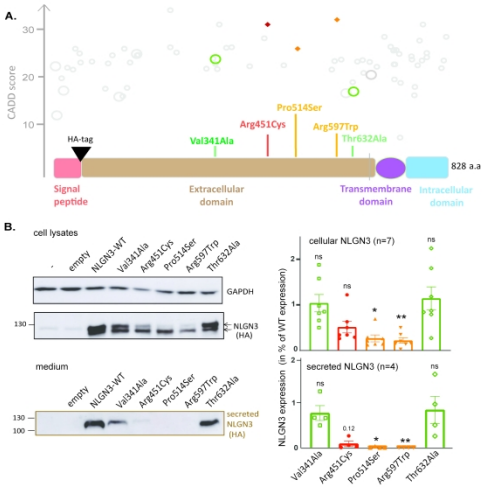
Supplementary Tables

Supplementary Table 1: List of nonsynonymous variants reported in *NLGN3* in the “non-neuro” cohorts from the gnomAD database.

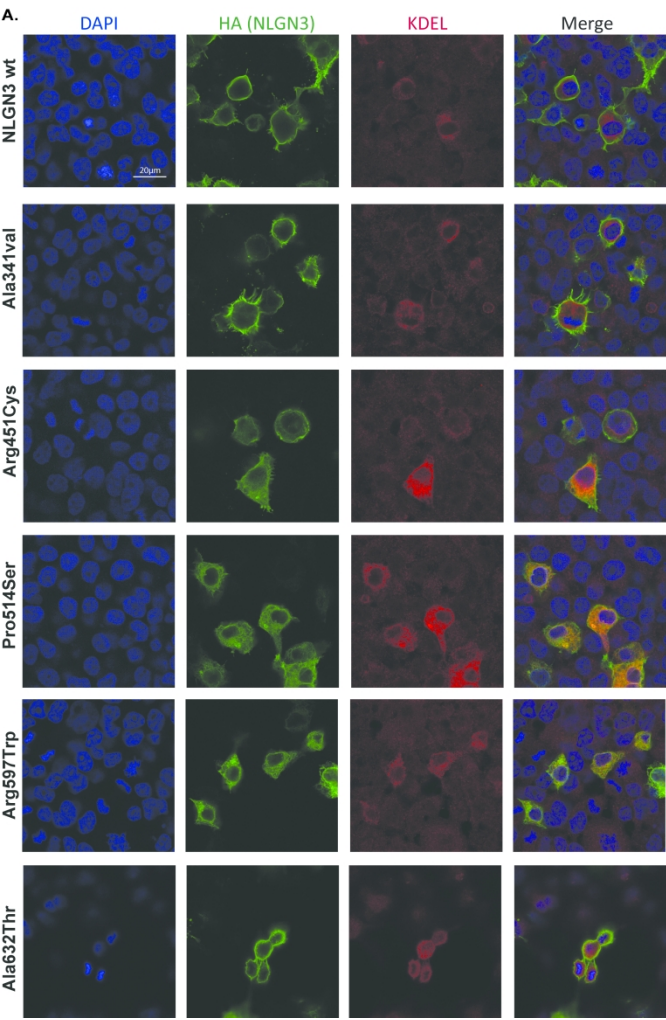
Supplementary Table 2: List of variants reported in *NLGN3* in the ClinVar database



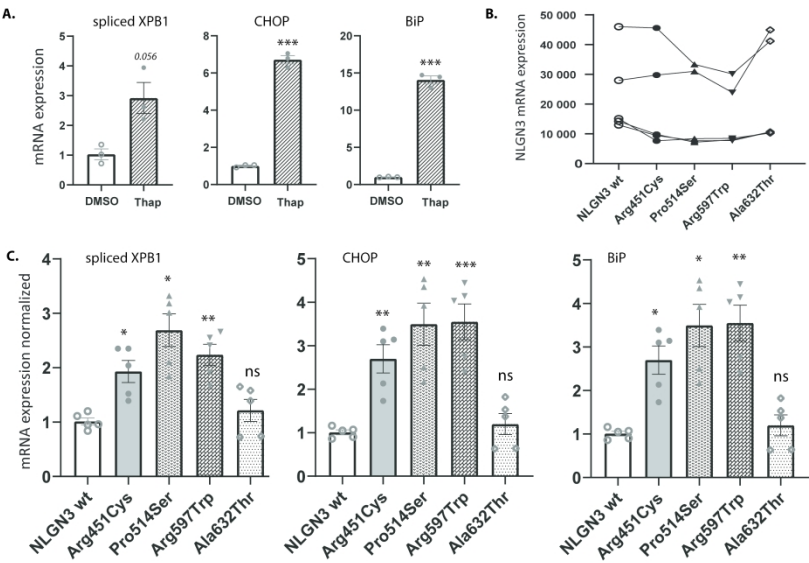
329x292mm (300 x 300 DPI)



465x282mm (300 x 300 DPI)



320x322mm (300 x 300 DPI)



342x198mm (300 x 300 DPI)

Table 1. Number of individuals with ASD previously screened for mutations in *NLGN3*

Publication	Year	N =	Population
Vincent et al.	2004	196	Canadian
Ylisaukko-oja et al.	2005	30	Finnish
Gauthier et al.	2005	96	Canadian
Blasi et al.	2006	124	IMGSAC
Talebizadeh et al.	2006	10	American
Wermter et al.	2008	107	Germany
Pampanos et al.	2009	169	Grece
Yanagi et al.	2012	62	Japanese
Avdjieva-Tzavella et al.	2012	20	Bulgarian
Steinberg et al.	2012	144	AGRE
Volaki et al.	2013	40	Greek
Xu et al.	2014	318*	Chinese
Mikhailov et al.	2014	143	Thai
Total		1459	

IMGSAC: International Molecular Genetic Study of Autism Consortium; N: number of individuals with ASD screened. *: a missense variant c.1276G>A, Gly406Ser (Gly426Ser according to NM_181303.1) was identified.

Nomenclature (NM_018977.3) <i>according to NM_181303.1</i>	Arg451Cys chrX:70387358C>T c.1351C>T c.1411C>T, p.Arg471Cys	Pro514Ser chrX:70387547C>T c.1540C>T c.1600C>T; p.Pro534Ser	Arg597Trp chrX:70389249C>T c.1789C>T c.1849C>T, p.Arg617Trp	Val341Ala chrX:70387029T>C c.1022T>C c.1082T>C, p.Val361Ala	Thr632Ala chrX:70389354A>G c.1894A>G c.1954A>G, p.Thr652Ala
Reference dbSNP	rs121917893	-	rs878853147	rs749067360	rs144914894
GnomAD^a	0	0	0	7	138
Grantham	180	74	101	64	58
SIFT	deleterious (0)	deleterious (0,02)	deleterious (0)	deleterious (0,05)	tolerated (0,49)
PolyPhen2	probably_damaging (1)	probably_damaging (0,97)	probably_damaging (1)	benign (0,13)	benign (0,00)
SNAP2	effect (37)	effect (40)	effect (67)	neutral (-23)	neutral (-51)
MutPred2	OI, TP, U	OI, TP, S	OI, RSA, DB, TP	TP	-
CADD	31	25.9	32	23.7	16.89
REVEL	0.767	0.810	0.769	0.437	0.125

Table 2. Summary of the missense variants tested. *In silico* predictions of functional consequences of the different missense changes tested in this study, assessed using several bioinformatics programs including SIFT, PolyPhen2, SNAP2 and MutPred2, CADD and REVEL (the two last ones compile information from the other programs). MutPred2 predictions include: “OI”: altered ordered interface; “TP”: altered transmembrane protein; “U”: loss of ubiquitylation; “S”: loss of sulfation; “RSA”: loss of relative solvent accessibility; “DB”: altered DNA Binding. ^a: Number of males from the “non-neuro” gnomAD cohorts who are hemizygous for the variant.